Significance of a decision making problem under uncertainty

Kevin Loquin
LIRMM
161 rue Ada
34095 Montpellier Cedex 5 France
kevin.loquin@lirmm.fr

Abstract

In this paper, we work on the interval dominance based extension of the Savage Expected Utility Maximization (SEUM) approach. While usual probabilities only handle variability due uncertainty, imprecise probabilities additionally handle, in a unique framework, epistemic uncertainty. This side of uncertainty, often called imprecision, can generate incomparability between the acts of a decision problem. Incomparability is linked to information held by the imprecise probability model quantifying the outcomes uncertainty. Our proposal, in this paper, is that for a given decision problem, its significance is the quantity of information which makes the interval dominance based imprecise SEUM decision problem change from incomparable to decidable (and possibly still not comparable) or comparable (and possibly still not decidable). We discuss incomparability sources, a theoretical and a pragmatical definition of significance of a decision problem under uncertainty.

Keywords. Savage EUM, decision theory, imprecise probability, interval dominance, significance

1 Introduction

Decision making boils down to comparing the outcomes of many possible acts. Modeling a decision problem (DP) is as simple as ranking the set of possible acts according to a preference relation generally constructed from a quantification of the consequences of each act (by means of a utility function). The distinction between the notions of comparability and decidability is very important in our work. A DP is said to be comparable when its set of acts can be ranked according to a complete preference relation. A DP is said to be decidable when there is a unique act which is optimal according to its preference relation.

When no uncertainty pertains the problem, the preference relation is naturally complete: any act can be ranked according to this preference relation. Even if the DP is not decidable, optimal choice(s) can always be found.

Decision making under uncertainty stands for situations when an act does not lead to a unique outcome with certainty. Since the preference relation between the acts is constructed from their outcomes, it seems natural to admit incomparability when facing uncertainty. Nevertheless, what is important for most decision maker is to work with a comparable DP and if possible with a decidable one. Most of the last century advances in decision making under uncertainty aimed at making complete, the ranking between the acts, by means of axioms which are supposed to be consistent with a rational (subjective) behavior. Among them, the Savage axiomatic, from which is derived the Expected Utility Maximization criterion (SEUM), is the most popular one [14].

Our view is that it is an artificial and arbitrary task to force the comparability of a DP under uncertainty. Indeed, due to partial information, a DP is inherently incomparable. We propose to ground our definition of significance on this “informativist” view of decision under uncertainty. Thus, we propose to work with imprecise probability based decisions theories [11, 12, 17, 18] instead of usual probability based decision theories [14]. The main asset of imprecise probability theories [1, 16, 18] over the usual probability theory is that they handle partial information. Within these schemes, uncertainty is characterized by interval weights (probability or more generally utility) instead of point-valued weights due to partial information. An imprecise probability model is a convex family of precise probabilities.

As is done in SEUM, some of the imprecise probability based decision theories propose a complete ranking between acts (maximin, maximax or Hurwicz criteria [11]), but what is the main richness of the imprecise probability based decision models is that they can admit incomparability due to lack of knowledge in a rigorous way. For instance, the E-admissibility criterion of Isaac Levi [12], the maximality as proposed by Peter Walley [18] or the very simple interval dominance decision rule all admit incomplete preference relations between the acts.

We aim at proposing a notion of significance of a DP based on this informativity interpretation of uncertainty in decision theory. Let us take any DP under uncertainty (thus incomparable or undecidable in the general case), its significance is the smallest quantity of information required to make it comparable or decidable. This is a quite intuitive idea: one faces a decision problem which is incomparable, the amount of required information to disambiguate the problem naturally characterizes the significance of the original problem. In some sense, significance aims at measuring the missing information for making the DP complete or decidable.

This definition of significance of a DP under uncertainty is abstract and is not grounded on any decision or uncertainty theory. In order to derive more concrete definitions, we propose to define the significance of a DP under uncertainty from the interval dominance decision rule and imprecise probability assessments over the outcome space.

Section 2 is a reminder (or a presentation) about the lower prevision model which summarizes the materials required for a proper understanding of this paper. The notion of imprecise expectation is particularly stressed. Section 3 presents the usual generalizations to imprecise probabilities of the SEUM. Finally Section 4 discuss the notion of significance of a DP under uncertainty as we aim at presenting it. A general (unrealistic) definition of significance is proposed, followed by a pragmatic definition of a significance index of a DP. A toy example inspired from [8] is also proposed to illustrate the notion of significance.

2 Imprecise Probability Theory

It is generally obvious (and probably non discussable) for most readers that the uncertainty about the outcome of any experiment is modeled by a set of precise weights between 0 and 1 on all the possible outcomes of this experiment: the probability weights. The general idea behind most of the imprecise probability theories is that uncertainty should preferably be modeled by a set of probability weights in order to handle imprecision, partial information or lack of knowledge inherent to most systems. Such new uncertainty theories are more general and powerful models than probability because they jointly and consistently handle the distinct notions of uncertainty due to variability and uncertainty due to imprecision that is generally called epistemic uncertainty.

2.1 Imprecise Probability Models

This theory presentation (which can be bypassed by expert readers) will emphasize on lower previsions defined on *discrete* domains, i.e. on domains with finite cardinality.

Let X be an uncertain variable whose possible outcomes are on a (finite) space $\mathcal{X}$ containing N exclusive single elements. Let $\mathcal{L}(\mathcal{X})$ denote the set of bounded real-valued functions on $\mathcal{X}$. $\mathcal{L}(\mathcal{X})$ is called the set of gambles. Each element (gamble) $f \in \mathcal{L}(\mathcal{X})$ is interpreted as the function on $\mathcal{X}$ representing the rewards $f(x)$ associated to the occurence of any possible outcome $x \in \mathcal{X}$ of X. Since the outcome value $x \in \mathcal{X}$ is uncertain, $f(x)$ is also an uncertain reward and thus f is an uncertain gamble.

A lower prevision $\underline{E}$ on $\mathcal{L}(\mathcal{X})$ is defined as a mapping $\underline{E} : \mathcal{K} \subseteq \mathcal{L}(\mathcal{X}) \to \mathbb{R}$. Its behavioral interpretation advocated by Walley is as follows: $\underline{E}(f)$ is interpreted as the supremum buying price an agent would accept for the uncertain reward $f(x)$. In order to ease the understanding of this fundamental concept, a lower prevision $\underline{E}(f)$ can be seen as the lower bound of the expectations of the uncertain gamble f. To a lower prevision $\underline{E}$ is associated its dual upper prevision $\overline{E}$ (or upper expectation), defined as $\overline{E}(f) = -\underline{E}(-f)$.

A lower prevision is said to be **coherent** on its gamble domain $\mathcal{K} \subseteq \mathcal{L}(\mathcal{X})$ if it satisfies the following conditions:

(C1) $\underline{E}(f) \geq \inf_{x\in\mathcal{X}} f(x)$ for all $f \in \mathcal{K}$ (accepting sure gain);

(C2) $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for each $f \in \mathcal{K}$ and $\lambda \geq 0$ (positive homogeneity);

(C3) $\underline{E}(f+g) \geq \underline{E}(f) + \underline{E}(g)$ for all $f, g \in \mathcal{K}$ (super-additivity).

A less restrictive class of lower previsions is the class of lower previsions avoiding sure loss. Let $\overline{G}(f)$ be the highest expected gain with a gamble $f \in \mathcal{K}$. It is naturally defined by $\overline{G}(f) = f - \underline{E}(f)$. We thus have a loss on f for the assessed prevision $\underline{E}$ when $\overline{G}(f) < 0$. Thus a lower prevision model is said **avoiding sure loss** when there is a set of gambles $(f_j)_{j=1,\dots,n}$ of $\mathcal{K}$ fulfilling $\sum_{i=1}^{n} \overline{G}(f_j) \geq 0$, i.e. when there is at least a set of gambles whose combination avoids a sure loss.

A coherent lower prevision which is equal to its associated upper prevision is said to be linear. Therefore, a coherent linear prevision denoted by E, i.e. such that for all $f \in \mathcal{K}$, $\underline{E}(f) = \overline{E}(f)$ fulfills both super-additivity **(C3)** and sub-additivity[1] and thus the finite additivity axiom: $E(f+g) = E(f) + E(g)$. A linear prevision can be seen as a usual expectation operator.

A lower prevision can be associated to a convex set of linear previsions. The set of linear previsions dominating the coherent lower prevision $\underline{E}$, defined on $\mathcal{K}$, called the credal set, is defined by:

$$\mathcal{M}(\underline{E}) = \{E \in \mathcal{E}(\mathcal{X}) \mid (\forall f \in \mathcal{K}) \, (\underline{E}(f) \leq E(f))\}, \quad (1)$$

where $\mathcal{E}(\mathcal{X})$ is the set of linear previsions on $\mathcal{X}$.

This object is particularly interesting since it links a lower expectation to its associated coherent set of dominating expectations.

An important particular case of coherent lower prevision is the lower probability. To any subset (or event) A of $\mathcal{X}$ can be associated its indicator function, which is a gamble, $\mathbb{1}_A \in \mathcal{L}(\mathcal{X})$. The lower probability of an event $A \subset \mathcal{X}$, denoted by $\underline{P}(A)$, is the lower prevision associated to this gamble $\mathbb{1}_A$. We denote by $\mathcal{B}(\mathcal{X})$, the set of indicator functions on $\mathcal{X}$, in order to remind the Borel algebra: $\mathcal{B}(\mathcal{X}) \subset \mathcal{L}(\mathcal{X})$ can be seen as the set of events on $\mathcal{X}$. To a lower probability is associated the dual notion of upper probability $\overline{P}(A) = 1 - \underline{P}(A^c)$, where A^c denotes the complement of A on $\mathcal{X}$.

Many other particular cases of the lower prevision model exist [3, 4, 19] that match the following inclusion: a necessity measure (dual of a possibility measure) is a particular case of belief function (whose pieces of evidence are consonant or nested [6]) ; a belief function is a particular case of convex Choquet Capacity ; a convex Choquet capacity is a particular case of lower probability ; a lower probability is a particular case of lower prevision.

2.2 Imprecise Expectation and natural extension

Facing a given quantity of information, a complete modeling of an uncertain variable X is done when a coherent lower prevision can be associated to any possible gamble f of $\mathcal{L}(\mathcal{X})$. However, in most real applications, information is limited to a lower prevision, denoted by $\underline{E}_{\mathcal{K}}$, defined on a subset of gambles $\mathcal{K} \subset \mathcal{L}(\mathcal{X})$. The natural extension procedure allows distributing (conveying) information held by $\underline{E}_{\mathcal{K}}$ to $\mathcal{L}(\mathcal{X})$ in the most conservative way. In other words, the natural extension is the most specific model, denoted by $\underline{E}$, that can be constructed on $\mathcal{L}(\mathcal{X})$ without any additional information incorporation, i.e. without reducing the model $\underline{E}_{\mathcal{K}}$.

[1]Sub-additivity: axiom **(C3)** with the reverse inequality.

Definition 2.1 (Natural Extension) *Suppose $\underline{E}_{\mathcal{K}}$ is a lower prevision on $\mathcal{K} \subset \mathcal{L}(\mathcal{X})$, then its natural extension $\underline{E}$ is defined, for any $f \in \mathcal{L}(\mathcal{X})$, by*

$$\underline{E}(f) = \sup_{\mathbb{R}} \left\{ \begin{array}{l} \alpha : f - \alpha \geq \sum_{j=1}^{n} \lambda_j (f_j - \underline{E}_{\mathcal{K}}(f_j)), \\ \textit{for some } n \geq 0, f_j \in \mathcal{K}, \lambda_j \geq 0 \end{array} \right\}. \quad (2)$$

$\underline{E}(f)$ is the supremum buying price for the gamble f given that the linear combination of the highest gain $\overline{G}(f_j) = f_j - \underline{E}_{\mathcal{K}}(f_j)$ associated to any set of gambles f_j of $\mathcal{K}$ is still higher than the gain $\overline{G}(f)$ obtained on the gamble f for this price $\underline{E}(f)$.

When $\underline{E}_{\mathcal{K}}$ avoids sure loss $\underline{E}$ is the minimal coherent lower prevision which dominates $\underline{E}_{\mathcal{K}}$ on $\mathcal{K}$. This gives all its meaning to the expression "in the most conservative way", which characterizes the way an uncertainty model $\underline{E}_{\mathcal{K}}$ on $\mathcal{K}$ is extended to $\underline{E}$ on $\mathcal{L}(\Omega)$. Note also that when $\underline{E}_{\mathcal{K}}$ is coherent, $\underline{E}_{\mathcal{K}}$ and $\underline{E}$ coincide on $\mathcal{K}$.

This tool is of prime importance since it tells us how to accomplish inference from the assessment of an imprecise prevision model on $\mathcal{K} \subset \mathcal{L}(\mathcal{X})$ to any gamble of $\mathcal{L}(\mathcal{X})$.

An interesting particular case is when $\mathcal{K} = \mathcal{B}(\mathcal{X})$. In that case, the natural extension procedure coincides with the computation of the lower expectation of any bounded function f: $\underline{E}(f)$ associated to the constraints provided by the lower probability model $\underline{P}$ defined on $\mathcal{B}(\mathcal{X})$. This is exactly what defines the imprecise expectation: this is the natural extension to $\mathcal{L}(\mathcal{X})$ of a lower prevision defined on $\mathcal{B}(\mathcal{X})$ (thus of a lower probability).

The imprecise expectation can only, in the general case of lower probability, be computed by using linear programming techniques. But, for a convex capacity (and any of its submodels: necessity, belief function,...), imprecise expectation can be computed by means of the Choquet integral [2].

3 Decision under uncertainty with Imprecise Probability

Uncertainty modeling has many available distinct theories generally associated to different interpretations. Decision modeling under uncertainty shares the same kind of diversity in its theories. In this section, we present the most encountered decision theories under uncertainty.

3.1 SEUM decision theory

In the SEUM decision theory, there is the set of possible acts, denoted by $\mathcal{A}$. Decision making under uncertainty stands for situations when an act does not lead, in general, to a unique outcome with certainty. Each act X of $\mathcal{A}$ is an uncertain variable with value in the finite outcome space, denoted by $\mathcal{X}$. This outcome space is a rather abstract space which can be numerical or not. For instance *patient healing* or *flood* are non numerical outcomes encountered in usual DP under uncertainty in the fields of medical decision or environmental risk assessment. In the SEUM approach, a utility function on the outcome space is used: $u : \mathcal{X} \to \mathbb{R}$ to quantify (and possibly rank) the acts (or their outcomes) on a utility scale.

Under the Savage axioms, the following complete preference relation $\succeq$ is constructed on $\mathcal{A}$ and defined, for X and Y in $\mathcal{A}$ by:

$$X \succeq Y \text{ iff } E_X(u) \geq E_Y(u). \tag{3}$$

In other words, an act X is preferred to another act Y when its associated expected utility is higher than the expected utility associated to Y. The optimal act(s) X^* is (are) such that

$$X^* \succeq Y,\ \forall Y \in \mathcal{A}.$$

At this point it is interesting to link some notations of the SEUM approach to notations of our imprecise probability presentation (IP) of Section 2. For instance, a gamble f in IP theory is similar to the utility function u of SEUM. Besides, the uncertain variable of IP and the uncertain outcome of SEUM, both denoted by X, are similar objects. We chose to incorporate the uncertain variable X to our IP presentation, which is generally not present in Walley theory and especially not in Walley's book [18], since it can easily be linked to the uncertain outcome of usual decision theories under uncertainty.

SEUM is a very elegant axiomatic construction [14] which entails a rational interpretation to preference structure (3). Many authors discussed and criticized the foundations of this approach by stressing too strong axioms [10]. Perhaps the most severe and constructive criticism is due to Ellsberg [7]. The SEUM is based on the idea that a decision maker behaves as if he possesses a complete and exhaustive knowledge of the possible states of the world, and moreover that, his assessment of the uncertainty about the outcomes may be represented as a unique finitely additive probability model. This idea has been termed as *probabilistic sophistication* [13]. Experimental evidence, as the Ellsberg paradox [7], has failed to support probabilistic sophistication as a good descriptive theory of behavior under uncertainty.

3.2 Imprecise SEUM generalizations and associated decision rules

Questioning the probabilistic sophistication principle of the SEUM approach has been done for many submodels of the lower prevision model: for possibility theory [5], for belief functions [10] or for capacities [15]. In such particular cases, the usual expectation operator based on the Lebesgue integral is replaced by a two-fold Choquet integral to compute the bounds of an imprecise utility expectation operator. The most general framework, i.e. obtained when uncertainty about the outcomes is modeled by a lower prevision $\underline{P}$, is computationally less tractable since it does not involve an explicit formulation of the imprecise expectation bounds but only linear optimization techniques.

Actually, most proposed generalizations of the SEUM to lower previsions were reduced to find meaningful ways to compare imprecise quantities: the imprecise expected utilities $[\underline{E}_X(u), \overline{E}_X(u)]$ instead of comparing precise quantities: the expected utilities $E_X(u)$. In other words, most approaches aim at finding a meaningful way to fulfill the first Savage axiom (which claims that *a preference relation is a complete ordering on the set of possible acts* $\mathcal{A}$) when the compared quantities are imprecise.

In order to expose some of the most encountered approaches, it is interesting to provide interpretations to $[\underline{E}_X(u), \overline{E}_X(u)]$. If u is a utility function on $\mathcal{X}$, $u(x)$ is uncertain due to the uncertainty on the outcomes of the act X, thus $\underline{E}_X(u)$ can be considered as the pessimistic expected utility associated to act X and $\overline{E}_X(u)$ can be considered as the optimistic expected utility associated to act X. In such framework, $u(x)$ represents a reward. It is therefore intuitive to term as optimistic the highest reward we can expect for uncertain outcomes of act X, i.e. $\overline{E}_X(u)$. Conversely, being pessimistic is to consider only the lowest reward we can expect with such model, i.e. $\underline{E}_X(u)$. Another remark is that when we are optimistic on a reward, we are pessimistic on a loss (and conversely) which is translated by relation $\overline{E}_X(u) = -\underline{E}_X(-u)$, since $-u$ is a loss when u is a reward.

This relevant interpretations of $\underline{E}_X(u)$ and $\overline{E}_X(u)$ as respectively the pessimistic and optimistic expected utility lead to propose a parametric optimal decision rule: the Hurwicz criterion, whose parameter r is a marker of the risk aversion of the decision maker. Actually, a decision maker has a high level of risk aversion when he considers for comparative quantities in its DP, the pessimistic expected utility. To favor the less risky problem posing and to be pessimistic are equivalent. Thus under a risk aversion (or pessimistic) attitude, the optimal decision rule is given by

$$X \succeq_P Y \text{ iff } \underline{E}_X(u) \geq \underline{E}_Y(u). \quad (4)$$

Optimism and risk are generally in accordance, thus the optimistic optimal decision rule is

$$X \succeq_O Y \text{ iff } \overline{E}_X(u) \geq \overline{E}_Y(u). \quad (5)$$

As a tradeoff between these rules stands the Hurwicz criterion. It is based on defining the expected utility for a risk aversion degree of r by :

$$E^r_X(u) = r\underline{E}_X(u) + (1-r)\overline{E}_X(u). \quad (6)$$

r is a sensible risk aversion index since $E^1_X(u) = \underline{E}_X(u)$ and $E^0_X(u) = \overline{E}_X(u)$. Thus the Hurwicz decision rule for a risk aversion degree r is

$$X \succeq^r_H Y \text{ iff } E^r_X(u) \geq E^r_Y(u). \quad (7)$$

Note that $\succeq_P$ is exactly $\succeq^1_H$ and $\succeq_O$ is exactly $\succeq^0_H$.

While the imprecise probability framework is supposed to model imprecision or epistemic uncertainty, to our view, the only consistent approaches, regarding this "informativist" view, are the approaches which allow incomparability between acts. At first sight admitting incomparability is problematic for providing optimal choices. However, this is a quite intuitive idea when facing epistemic uncertainty. In most cases where information is partial, admitting incomparability (and/or indecision) is safer than proposing a choice even if this choice is supposed to be obtained with a pessimistic rule. Let us consider an example of cancer diagnosis which illustrates a rational behavior under epistemic uncertainty: for most kind of cancers, abnormal blood tests results are not significant enough to diagnose cancer and an additional biopsy is generally required. Thus when information is partial (only the blood tests result), the physician admits incomparability and thus indecision. He will never claim that the patient has cancer and decide to start a heavy chemotherapy treatment only from these partial evidences.

Three decision rules admitting incomparability between acts are generally considered, the E-admissibility of Isaac Levi [12], the maximality as proposed by Peter Walley [18] or the very simple interval dominance decision rule. Interval dominance criterion is defined through the following incomplete preference relation:

$$X \succeq_{ID} Y \text{ iff } \underline{E}_X(u) \geq \overline{E}_Y(u). \quad (8)$$

This is certainly the most intuitive and simple decision rule admitting incomparability with imprecise probability. It says that an act X is preferred to an act Y if the imprecise expected utility of X completely (in terms of interval) dominates the imprecise expected utility of Y.

Actually this is the most cautious rule. Indeed a DP which is not comparable for the interval dominance criterion can be comparable for the E-admissibility and/or the maximality criteria. It implicitly means that available information is considered as insufficient for the interval dominance criterion while sufficient for the other criteria.

3.3 Sources of incomparability: a discussion

As already mentioned, a DP is said to be comparable when its set of acts can be ranked according to a complete preference relation and a DP is said to be decidable when there is a unique act which is optimal according to its preference relation. There is no inclusion relation between the decidability and the comparability of a DP. A decidable problem is not necessarily comparable. This is the case if there exists an optimal act for a partial preference ordering. Conversely, a comparable problem is not necessarily decidable. This is the case for any problem which results in more than one indifferent optimal acts for a complete preference ordering.

In this paper, we propose to use the non comparability of a DP under uncertainty to define its significance. Thus, it is interesting to discuss the incomparability sources of an imprecise SEUM problem. To our view, the sources of incomparability are twofold: 1/ epistemic (or reducible) uncertainty but also 2/ the problem construction itself. While they may not be the only sources of incomparability of an imprecise SEUM problem, they are certainly among these sources.

Indeed, 1/ the influence of the epistemic uncertainty on the comparability of a DP can easily be shown: let us take any incomparable imprecise SEUM problem, if uncertainty is reduced to a precise probability model then we recover a usual (i.e. precise) SEUM and thus a comparable DP.

And, 2/ the influence of the problem construction itself can be put forward: let us consider two different problems (i.e. two different utility functions) but with the same set of acts and associated uncertain outcomes and the same imprecise probability assessments for these parameters. We denote (P1) and (P2) these imprecise SEUM problems. We can find cases where (P1) provides a comparable decision framework, while (P2) is still incomparable.

Among the other possible sources of incomparability, we were wondering if the imprecise expectation operator which is used to pass from the uncertainty assessment step to the comparison step of an imprecise SEUM problem, has some impact on the comparability of the problem. Our answer is not clear yet but we showed some continuity results of the imprecise expectation operator in a working paper. These results tend to prove that the imprecise expectation operator does not impact the comparability of the problem. Indeed, continuity means that variations (measured with Hausdorff distances) between imprecise expectations are bounded by the variations between their generative imprecise probability models. Such stability is important in imprecise SEUM. It means that information rooting the uncertainty assessment of an imprecise SEUM problem is properly conveyed to the utility comparison step. More than this topological stability, it was already said that the natural extension is the most conservative extension of an imprecise probability model to the expectation of a utility function (or gamble).

4 Significance of a decision making problem under uncertainty

Now, let us reexamine an already considered situation: we are facing two different problems (i.e. with two different utility functions) with the same uncertain outcomes. Let us consider that both problems are non comparable and non decidable. If we progressively reduce the epistemic uncertainty associated to the uncertain outcomes of the problems, one problem, for instance (P1), should become comparable or decidable before the other problem (P2). It is thus natural to claim that problem (P1) is more significant than problem (P2) regarding the original pieces of information. Indeed, (P1) requires less artificial information addition than (P2) to become decidable or comparable.

The previous paragraph is the heart of this paper, since it explains the notion of significance as we hear it. We will say that a DP under uncertainty is fully significant if its associated ranking of the set of acts $\mathcal{A}$ is complete for the interval dominance or is decidable (even if non comparable). A DP under uncertainty is fully insignificant when the system must be reduced to a precise SEUM to become a comparable DP (decidable or not). Between these extreme cases, we will define the significance index of an incomparable and undecidable DP: it is the smallest quantity of information required to make it comparable or decidable.

In a sense, significance, as we aim at defining it, is a measure of "missing information" to make the problem comparable or decidable. Thus significance is a measure of meta-information: information about information. As for imprecise SEUM problems, information is modeled by lower previsions. It models information about a true underlying probability measure. Thus, meta-information can only be consistently quantified if we know the true underlying probability. In other words, it is impossible to judge information (i.e. to quantify meta-information) without knowing the truth. That is the reason why we ground our first definition of significance on the (unrealistic and unapplicable) assumption that we know the true underlying probability of an imprecise SEUM problem.

Note that all the involved lower probabilities in this definition of significance are consistant with the true underlying probability. It means that we only work with information which are not conflicting. Thus, we do not compete with formal decision frameworks which deal with ambiguity and conflict as separate types of uncertainty [9].

4.1 An unrealistic general definition of a significance index

The most general (but unrealistic) definition of a significance index that we will propose requires some preliminary definitions and notations.

Let $\underline{P_X}$ be a lower probability on the act X, which is an uncertain variable with values in the outcome space $\mathcal{X}$. Let P_0 be the true underlying probability modeling the uncertainty about X. We assume that $\underline{P_X}$ is consistant with P_0, i.e. $P_0 \geq \underline{P_X}$.

Let $\underline{\mathcal{P}}(\underline{P_X}) = \{\underline{P} : P_0 \geq \underline{P} \geq \underline{P_X}\}$ be the set of lower probabilities consistent with P_0 and dominating $\underline{P_X}$. It is the set of lower probability models more specific than $\underline{P_X}$, i.e. more informed, and still consistant with $\underline{P_X}$.

Let d be a distance between imprecise probabilities of $\underline{\mathcal{P}}(\underline{P_X})$ which respects the domination. We mean that, for three encapsulated (according to heir specificity) lower probabilities $\underline{P_1}$, $\underline{P_2}$ and $\underline{P_3}$, such that $\underline{P_1} \leq \underline{P_2} \leq \underline{P_3}$ then $d(\underline{P_1}, \underline{P_2}) \leq d(\underline{P_1}, \underline{P_3})$. This property is quite natural since it enables to use such distance for ranking the lower probabilities specificity-wise relative to a given lower probability. For instance, $d(\underline{P_1}, \underline{P_2}) \leq d(\underline{P_1}, \underline{P_3})$ means that, relatively to $\underline{P_1}$, we have that $\underline{P_2} \leq \underline{P_3}$, i.e. that $\underline{P_2}$ is more specific than $\underline{P_3}$. Note that the Hausdorff distance between sets of probabilities and thus between lower probabilities fulfills such natural property. It should be interesting to study other distances between lower probabilities respecting this property.

Let d_0 be this distance between any lower probability $\underline{P}$ of $\underline{\mathcal{P}}(\underline{P_X})$ and P_0: $d_0(\underline{P}) = d(P_0, \underline{P})$. We also define

$$d_{0X} = d_0(\underline{P_X}) = d(P_0, \underline{P_X})$$

as the distance between $\underline{P_X}$ and P_0.

Let α be the distance between any lower probability $\underline{P} \in \underline{\mathcal{P}}(\underline{P_X})$ and P_0 relative to the distance between $\underline{P_X}$ and P_0. α is defined by

$$\alpha(\underline{P}) = \frac{d_0(\underline{P})}{d_{0X}}.$$

This relative distance is such that $\alpha(\underline{P}) \in [0, 1]$ for any lower probability $\underline{P} \in \underline{\mathcal{P}}(\underline{P_X})$ and $\alpha(P_0) = 0$ and $\alpha(\underline{P_X}) = 1$.

In other words, should we assume that P_0 exists and is known (which is not consistent with the Walley's behavioral imprecise probability framework), $\alpha(\underline{P})$ can be considered as a normalized index of non specificity (of imprecision) of $\underline{P}$.

Now, let us define, for a given imprecise SEUM problem (P), $\mathcal{C}$: the set of lower probabilities of $\underline{\mathcal{P}}(\underline{P_X})$ which make the problem comparable or decidable. Now we can propose a general unrealistic definition of the significance of an imprecise SEUM.

Definition 4.1 (Significance)
Let (P) be an imprecise SEUM problem: $\underline{P_X}$ is a lower probability on X defined on $\mathcal{X}$ and u is a utility function on $\mathcal{X}$.

Let $\underline{P}^*$, be the least specific lower probability of $\underline{\mathcal{P}}(\underline{P_X})$, which makes (P) comparable or decidable. Then the significance of (P) is given by

$$\mathcal{S}_{(P)} = \alpha(\underline{P}^*). \tag{9}$$

An alternative definition can be proposed:

$$\mathcal{S}_{(P)} = \max_{\underline{P} \in \mathcal{C}} \alpha(\underline{P}). \tag{10}$$

The interpretation we can propose to this index is as follows. Significance is the maximal degree of imprecision (of epistemic uncertainty) which allows comparability. For a lower prevision model with an imprecision higher than $\mathcal{S}_{(P)}$, the problem is still incomparable, but for a lower prevision model with an imprecision lower than $\mathcal{S}_{(P)}$, the problem i scomparable or decidable.

Let us retake the example presented in the first paragraph of Section 4.1. We can rephrase it that way: the highest imprecision which makes the problem comparable or decidable is bigger for (P1) than for (P2) thus $\mathcal{S}_{(P1)} \geq \mathcal{S}_{(P2)}$.

Finally, if we are facing a problem (P) which is comparable regarding the provided information $\underline{P_X}$, then the significance of this problem should be the highest, i.e. should be equal to 1. With our definition, $\mathcal{S}_{(P)} = 1$, since $\mathcal{C} = \underline{\mathcal{P}}(\underline{P_X})$ and $\alpha(\underline{P_X}) = 1$. On the contrary, if we are facing a problem (P') which is comparable or decidable only when uncertainty is reduced to a linear probability, then the significance of this problem should be the lowest, i.e. should be equal to 0. With our definition, $\mathcal{S}_{(P')} = 0$, since $\mathcal{C} = \{P_0\}$ and $\alpha(P_0) = 0$.

4.2 Significance index : an applicable definition

Definition 4.1 of the significance is not applicable because P_0 is unknown (even if it exists). We propose in this section a pragmatic significance index for the imprecise SEUM approach with the interval dominance rule.

In Definition 4.1, the imprecision reduction is performed directly on the lower probability $\underline{P_X}$ modeling the uncertainty about X. In the applicable definition, we propose to perform this imprecision reduction directly on the interval utility expectations $[\underline{E}_X(u), \overline{E}_X(u)]$ associated to every act X.

This applicable definition is inspired from the Hurwicz risk aversion degree (6). In our case we define the relative imprecision index ρ of the imprecise expected utility as:

$$\begin{cases} \underline{E}^\rho_X(u) = (1-\rho)E_0(u) + \rho\underline{E}_X(u), \\ \overline{E}^\rho_X(u) = (1-\rho)E_0(u) + \rho\overline{E}_X(u), \end{cases} \tag{11}$$

where $E_0(u) = \frac{\underline{E}_X(u)+\overline{E}_X(u)}{2}$ is the middle of $[\underline{E}_X(u), \overline{E}_X(u)]$.

ρ is an index of imprecision relative to the imprecision of $\underline{E_X}$. We interpret $[\underline{E}^\rho_X(u), \overline{E}^\rho_X(u)]$ as the representation of $[\underline{E}_X(u), \overline{E}_X(u)]$ of relative imprecision ρ. Indeed, for a relative imprecision $\rho = 0$, $[\underline{E}^0_X(u), \overline{E}^0_X(u)] = \{E_0(u)\}$ and for a relative imprecision $\rho = 1$, $[\underline{E}^1_X(u), \overline{E}^1_X(u)] = [\underline{E}_X(u), \overline{E}_X(u)]$. In other words, $[\underline{E}^\rho_X(u), \overline{E}^\rho_X(u)]$ goes from $\{E_0(u)\}$ to $[\underline{E}_X(u), \overline{E}_X(u)]$ when ρ goes from 0 to 1.

We can thus define a new decision rule which is called the ρ-imprecise decision rule and which is the interval dominance decision applied to the ρ-imprecise interval $[\underline{E}^\rho_X(u), \overline{E}^\rho_X(u)]$:

$$X \succeq_\rho Y \text{ iff } \underline{E}^\rho_X(u) \geq \overline{E}^\rho_Y(u). \tag{12}$$

The proposed definition of the applicable significance is thus a direct application of Definition 4.1.

Definition 4.2 (Applicable Significance)
Let (P) be an imprecise SEUM problem: $\underline{P_X}$ is a lower probability on X defined on $\mathcal{X}$ and u is a utility

function on $\mathcal{X}$. Let ρ^*, be the highest relative imprecision index, such that $\succeq_{\rho^*}$ becomes complete or makes (P) decidable. Then

$$\mathcal{S}_{(P)} = \rho^*. \qquad (13)$$

Compared to Definition 4.1, this solution, Definition 4.2 is feasible. Anyway, artificially increasing the informativity of an imprecise probability model is the only possible way to propose an applicable significance index. Indeed the informativity of any model can only be measured if we know the underlying true model, which is impossible or artificially possible.

Now let us illustrate this notion of significance on a toy example taken from [8].

Example

Assume that an individual with initial wealth ω is facing a risk of loss ℓ. There is uncertainty about the fact that this loss occurs or not. Each act X has two possible rewards: one if loss occurs, denoted by x_ℓ, and one if loss does not occur, denoted by $x_{\bar{\ell}}$.

One possible act for the individual would be not to buy any insurance. This can be represented by the act $X = (x_\ell, x_{\bar{\ell}}) = (\omega - \ell, \omega)$. Another act would be to buy full coverage at a premium π, yielding $Y = (y_\ell, y_{\bar{\ell}}) = (\omega - \pi, \omega - \pi)$. A third possible act would be to buy partial coverage at a premium π', yielding $Z = (z_\ell, z_{\bar{\ell}}) = (\omega - \ell + I - \pi', \omega - \pi')$ where I is the indemnity paid in case of damage.

We assume that the individual wealth is $\omega = \frac{3}{2}$, that its potential loss $\ell = \frac{1}{2}$, that the respective full and partial coverage are given by $\pi = \frac{1}{5}$ and $\pi' = \frac{1}{10}$ and that the indemnity is $I = \frac{1}{3}$. We also assume that the imprecise probability of loss is given by $\{(p, 1-p) :$ for $p \in [\frac{1}{3}, \frac{1}{2}]\}$. The utility function is $u(x) = x$ for $x \in \mathcal{X}$. Under such assumptions, the compared imprecise expectations are given by:

- $[\underline{E}_X(u), \overline{E}_X(u)] = [1.25, 1.33]$,
- $[\underline{E}_Y(u), \overline{E}_Y(u)] = \{1.3\}$,
- $[\underline{E}_Z(u), \overline{E}_Z(u)] = [1.288, 1.3166]$.

We can compute easily that the significance of this DP is 0.2 and that the associated optimal decision is Z. Indeed, for decreasing relative imprecision indices, Table 1 shows the evolution of the imprecise utility expectation when we artificially decrease imprecision.

We can see from Table 1 that the DP becomes decidable and completely ranked for $\rho = 0.2$ and that the associated optimal choice is Z. In other words, with a significance of 0.2 the individual should choose to buy the proposed partial coverage π'.

ρ	$[\underline{\overline{E}}_X(u)]$	$[\underline{\overline{E}}_Y(u)]$	$[\underline{\overline{E}}_Z(u)]$
0.3	[1.278, 1.302]	1.3	[1.2986 , 1.3069]
0.2	[1,282 , 1,298]	1.3	[1.3, 1.3056]
0.1	[1.286, 1.294]	1.3	[1.3014, 1.3042]

Table 1: Imprecise utility expectations for various relative imprecision

End of Example

It should be noted that the aim of our proposal is not to provide an optimal decision. Actually, with Definition 4.2, the optimal choice(s) is (are) always the optimal choice(s) for the center of the utility expectation intervals associated to the acts. The Hurwicz criterion with a risk aversion of $r = \frac{1}{2}$ gives the same result, i.e. the same optimal choice(s). However, our approach aims at providing a significance index which is not done with the Hurwicz criterion or any other decision rule. The proposed simplified and pragmatic definition is a simple way to explain and introduce the notions of interest in this paper. But more sensible and complex definitions of significance should be proposed in later works.

5 Conclusion

This article is a discussion paper. Its aim is mainly to define a new notion of significance of decision problem under uncertainty and to discuss its foundations. The idea is that if a decision problem is not comparable then the quantity of information which is required to make it comparable or decidable is directly linked to its significance. A theoretical definition of a significance index is proposed. This definition is constructed with the true underlying model of an imprecise probability and is thus unrealistic. A second artificial but pragmatical index is proposed. This index is very simple and inspired from the way the Hurwizc decision criterion is constructed.

The next step is to derive explicit formulations of other significance indices based on pragmatic constructions similar or different than the one found in Section 4.2 and obtained for different imprecise probability models. For instance with any submodel of the convex Choquet capacities, the imprecise expectation is explicitly computed with the Choquet integral. Thus explicit formulations of significance indices are possible. Experimental studies are now to be proposed.

Acknowledgment

This work has been supported by INSERM through Plan Cancer Program 2009-2013

References

[1] Dempster, A.: Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics 38, 325–339 (1967)

[2] Denneberg, D.: Non Additive Measure and Integral. Kluwer Academic Publishers, Dordrecht (1994)

[3] Destercke, S., Dubois, D., Chojnacki, E.: Unifying practical uncertainty representations: I. generalized p-boxes. International Journal of Approximate Reasoning 49(3), 649–663 (2008)

[4] Destercke, S., Dubois, D., Chojnacki, E.: Unifying practical uncertainty representations: Ii. clouds. International Journal of Approximate Reasoning 49(3), 664–677 (2008)

[5] Dubois, D., Fargier, H., Perny, P.: Qualitative decision theory with preference relations and comparative uncertainty: An axiomatic approach. Artificial Intelligence 148(1-2), 219 – 260 (2003)

[6] Dubois, D., Prade, H.: Consonant approximations of belief functions. International Journal of Approximate Reasoning 4(5-6), 419–449 (1990)

[7] Ellsberg, D.: Risk, ambiguity, and the savage axioms. The Quarterly Journal of Economics 75, 643–669 (1961)

[8] Etner, J., Jeleva, M., Tallon, J.M.: Decision theory under ambiguity. Journal of Economic Surveys 26(2), 234–270 (2012)

[9] Gajdos, T., Vergnaud, J.C.: Decisions with conflicting and imprecise information. Social Choice and Welfare pp. 1–26 (2009)

[10] Jaffray, J., Wakker, P.: Decision making with belief functions: Compatibility and incompatibility with the sure-thing principle. Journal of Risk and Uncertainty 7, 255–271 (1993)

[11] Jaffray, J.: Utility theory for belief functions. Operations Research Letters 8, 107–112 (1989)

[12] Levi, I.: The Enterprise of Knowledge. MIT Press, Cambridge MA (1980)

[13] Marinacci, M.: Probabilistic sophistication and multiple priors. Information Science 1, 1–3 (2001)

[14] Savage, L.: The foundations of statistics. John Wiley & Sons (1954)

[15] Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrica 57, 571–587 (1989)

[16] Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)

[17] Troffaes, M.: Decision making under uncertainty using imprecise probabilities. International Journal of Approximate Reasoning 45(1), 17–29 (2007)

[18] Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, New York (1991)

[19] Walley, P.: Towards a unified theory of imprecise probability. International Journal of Approximate Reasoning 24(2-3), 125–148 (2000)